- - What is claimed is:

22. (New) A method for computer-aided determination of a sequence of actions for a system having states, the method comprising the steps of:

performing a transition in state between two states on the basis of an action;

determining the sequence of actions to be performed such that a sequence of states results from the sequence of actions;

optimizing the sequence of steps with regard to a prescribed optimization function, including a variable parameter; and

using the variable parameter to set a risk which the resulting sequence of states has with respect to a prescribed state of the system.

23. (New) The method as claimed in claim 22, further comprising the step of:

using approximative dynamic programming for the purpose of determination.

24. (New) The method as claimed in claim 23, further comprising the step of:

basing the approximative dynamic programming Q-learning.

25. (New) The method as claimed in claim 24, further comprising the steps of:

forming an optimization function within Q-learning in accordance with the following rule:

$$OFQ = Q\left(x;\ w^{a}\right),$$

and

adapting weights of the function approximator in accordance with the following rule:

$$w_{t+1}^{a_t} = w_t^{a_t} + \eta_t \cdot \aleph^{\kappa}\left(d_t\right) \cdot \nabla Q\left(x_t;\ w_t^{a_t}\right)$$

wherein

$$d_t = r\left(x_t, a_t, x_{t+1}\right) + \gamma \max_{a \in A} Q\left(x_{t+1}, w_t^{a}\right) - Q\left(x_t, w_t^{a_t}\right)$$

26. (New) The method as claimed in claim 23, further comprising the step of:

basing the approximative dynamic programming on TD(λ)-learning.

27. (New) The method as claimed in claim 26, further comprising the steps of:

forming the optimization function within TD(λ)-learning in accordance with the following rule:

OFTD = J(x;w); and

adapting weights of the function approximator are adapted in accordance with the following rule:

$w_{t+1} = w_t + \eta_t \cdot \aleph^{\kappa}(d_t) \cdot z_t$, wherein $d_t = r(w_t, a_t, x_{t+1}) + \gamma J(x_{t+1}; w_t) - J(x_t; w_t)$, $z_t = \lambda \cdot \gamma \cdot z_{t-1} + \nabla J(x_t; w_t)$, and $z_{-1} = 0$.

28. (New) The method as claimed in claim 27, further comprising the step of:

using a technical system to determine the sequence of actions before the determination measured values are measured.

29. (New) The method as claimed in claim 28, further comprising the step of:

subjecting the technical system to open-loop control in accordance with the sequence of actions.

cont

30. (New) The method as claimed in claim 28, further comprising the step of:

subjecting the technical system to closed-loop control in accordance with the sequence of actions.

31. (New) The method as claimed in claim 30, further comprising the step of:

modeling the system as a Markov Decision Problem.

32. (New) The method as claimed in claim 31, further comprising the step of:

using the system in a traffic management system.

33. (New) The method as claimed in claim 31, further comprising the step of:

using the system in a communications system.

a1 cont.

34. (New) The method as claimed in claim 31, further comprising the step of:

using the system to carry out access control in a communications network.

35. (New) The method as claimed in claim 31, further comprising the step of:

using the system to carry out routing in a communications network.

36. (New) A system for determining a sequence of actions for a system having states, wherein a transition in state between two states is performed on the basis of an action, the system comprising:

a processor for determining a sequence of actions, whereby a sequence of states resulting from the sequence of actions is optimized with regard to a prescribed optimization function, and the optimization function includes a variable parameter for setting a risk which the resulting sequence of states has with respect to a prescribed state of the system.

37. (New) The system as claimed in claim 36, wherein the processor is used to subject a technical system to open-loop control.

A1 Cont.

38. (New) The system as claimed in claim 36, wherein the processor is used to subject a technical system to closed-loop control. - -

39. The system as claimed in claim 36, wherein the processor is used in a traffic management system.